LOE	Method for Assessing LOE	Assessment Tools	Assessment(s)	Potential PRG Matrix	Strong Line? Why?	What issues affect reliability and certainty of LOE?	Can a Numeric PRG be Derived? (i.e., evidence of a relationship to sediment)	Corroborated by other LOEs? If so, which ones and what are their strengths.	LWG Position - Should this LOE be used to derive PRGs for use in FS?
Toxicity of sediment-associated chemicals to benthic invertebrates	Effect of laboratory exposure to Study Area sediment on *Chironomus dilutus* survival	10-day *Chironomus dilutus* survival test	Compare negative control-adjusted bioassay data to reference envelope values (REVs)	Sediment	Yes. Site-specific bioassay and correlates with sediment chemistry	Organism-level endpoint being used to assess community-level risk	Yes; in conjunction with a site-specific benthic toxicity model to tie back to sediment concentrations	Yes. 10-day *Chironomus dilutus* biomass test and 28-day *Hyalella azteca* survival test, benthic community data, surface water and tissue	Yes. The 10-day *Chironomus dilutus* survival test results and FPM should be used together to derive site-specific SQGs.
		Floating percentile model (FPM) for predicting 10-day *Chironomus dilutus* survival based on sediment chemical concentrations	Compare sediment chemical concentrations to FPM-derived site-specific sediment quality guidelines (SQGs)		Yes. FPM predicts empirically observed 10-day *Chironomus dilutus* survival test results	The FPM establishes correlations but not causation and doesn't provide a unique solution.	Yes, in conjunction with the 10-day *Chironomus dilutus* survival test results to tie back to sediment concentrations	Yes, 10-day *Chironomus dilutus* biomass FPM and 28-day *Hyalella azteca* survival FPM, benthic community data, surface water and tissue	
		Logistic regression model (LRM) for predicting 10-day *Chironomus dilutus* survival based on sediment chemical concentrations	Compare sediment chemical concentrations to LRM-derived site-specific SQGs		No. LRM doesn't predict empirically observed 10-day *Chironomus dilutus* survival test results	If it worked, the LRM would establish correlations but not causation and wouldn't provide a unique solution.	No	No	No. The LRM failed to reliably predict empirical bioassay hit classification results, whereas the FPM succeeded. The FPM and LRM are just two different ways to generate SQGs. Since one worked and the other didn't, the one that worked should be used and the one that didn't should be set aside.
	Effect of laboratory exposure to Study Area sediment on *Chironomus dilutus* biomass	10-day *Chironomus dilutus* biomass test	Compare negative control-adjusted bioassay data to REVs	Sediment	Yes. Site-specific bioassay and correlates with sediment chemistry	Organism-level endpoint being used to assess community-level risk	Yes; in conjunction with a site-specific benthic toxicity model to tie back to sediment concentrations	Yes. 10-day *Chironomus dilutus* and 28-day *Hyalella azteca* survival tests, benthic community data, surface water and tissue	Yes. The 10-day *Chironomus dilutus* biomass test results and FPM should be used together to derive site-specific SQGs.
		FPM for predicting 10-day *Chironomus dilutus* biomass based on sediment chemical concentrations	Compare sediment chemical concentrations to FPM-derived site-specific SQGs		Yes. FPM predicts empirically observed 10-day *Chironomus dilutus* biomass test results	The FPM establishes correlations but not causation and doesn't provide a unique solution.	Yes, in conjunction with the 10-day *Chironomus dilutus* biomass test results to tie back to sediment concentrations	Yes, 10-day *Chironomus dilutus* and 28-day *Hyalella azteca* survival FPMs, benthic community data, surface water and tissue	
		LRM for predicting 10-day *Chironomus dilutus* biomass based on sediment chemical concentrations	Compare sediment chemical concentrations to LRM-derived site-specific SQGs		No. LRM doesn't predict empirically observed 10-day *Chironomus dilutus* biomass test results	If it worked, the LRM would establish correlations but not causation and wouldn't provide a unique solution.	No	No	No. The LRM failed to reliably predict empirical bioassay hit classification results and the FPM succeeded. The FPM and LRM are just two different ways to generate SQGs. Since one worked and the other didn't the one that worked should be used and the one that didn't should be set aside.
	Effect of laboratory exposure to Study Area sediment on *Hyalella azteca* survival	28-day *Hyalella azteca* survival test	Compare negative control-adjusted bioassay data to REVs	Sediment	Yes. Site-specific bioassay and correlates with sediment chemistry	Organism-level endpoint being used to assess community-level risk	Yes; in conjunction with a site-specific benthic toxicity model to tie back to sediment concentrations	Yes. 10-day *Chironomus dilutus* survival and biomass tests, benthic community data, surface water and tissue	Yes. The 28-day *Hyalella azteca* survival test results and FPM should be used together to derive site-specific SQGs.

LOE	Method for Assessing LOE	Assessment Tools	Assessment(s)	Potential PRG Matrix	Strong Line? Why?	What issues affect reliability and certainty of LOE?	Can a Numeric PRG be Derived? (i.e., evidence of a relationship to sediment)	Corroborated by other LOEs? If so, which ones and what are their strengths.	LWG Position - Should this LOE be used to derive PRGs for use in FS?
		FPM for predicting 28-day *Hyalella azteca* survival based on sediment chemical concentrations	Compare sediment chemical concentrations to FPM-derived site-specific SQGs		Yes. FPM predicts empirically observed 28-day *Hyalella azteca* survival test results	The FPM establishes correlations but not causation and doesn't provide a unique solution.	Yes, in conjunction with the 28-day *Hyalella azteca* survival test results to tie back to sediment concentrations	Yes, 10-day *Chironomus dilutus* survival and biomass FPMs , benthic community data, surface water and tissue	
		LRM for predicting 28-day *Hyalella azteca* survival based on sediment chemical concentrations	Compare sediment chemical concentrations to LRM-derived site-specific SQGs		No. LRM doesn't predict empirically observed 28-day *Hyalella azteca* survival test results	If it worked, the LRM would establish correlations but not causation and wouldn't provide a unique solution.	No	No	No. The LRM failed to predict empirical bioassay hit classification results and the FPM succeeded. The FPM and LRM are just two different ways to generate SQGs. Since one worked and the other didn't the one that worked should be used and the one that didn't should be set aside.
	Effect of laboratory exposure to Study Area sediment on *Hyalella azteca* biomass	28-day *Hyalella azteca* biomass test	Compare negative control-adjusted bioassay data to REVs	Sediment	No. Site-specific bioassay but doesn't correlate with sediment chemistry	Organism-level endpoint being used to assess community-level risk	Yes; in conjunction with a site-specific benthic toxicity model to tie back to sediment concentrations (but they'll be unreliable because we've been unable to develop a reliable model)	No	No. The empirical bioassay data can only be used to derive PRGs (i.e., site-specific SQGs) in conjunction with a predictive model. The FPM and LRM both failed to predict hit classification results for the 28-day *Hyalella azteca* biomass endpoint. Reliable models were developed for the other three bioassay endpoints and they should be used instead of the *Hyalella* biomass endpoint to derive PRGs for use in the FS.
		FPM for predicting 28-day *Hyalella azteca* biomass based on sediment chemical concentrations	Compare sediment chemical concentrations to FPM-derived site-specific SQGs		No. FPM doesn't predict empirically observed 28-day *Hyalella azteca* biomass test results	The FPM establishes correlations but not causation and doesn't provide a unique solution.	Yes, in conjunction with the 28-day *Hyalella azteca* biomass test results to tie back to sediment concentrations (but they'll be unreliable because we've been unable to develop a reliable model)	No	
		LRM for predicting 28-day *Hyalella azteca* biomass based on sediment chemical concentrations	Compare sediment chemical concentrations to LRM-derived site-specific SQGs		No. LRM doesn't predict empirically observed 28-day *Hyalella azteca* biomass test results	If it worked, the LRM would establish correlations but not causation and wouldn't provide a unique solution.	No	No	
	Biological effects (broadly defined) of exposure to Study Area sediment on aquatic organisms (in general)	1. **Threshold effect levels (TELs)** for predicting freshwater sediment concentrations "rarely associated with adverse biological effects" (broadly defined) on aquatic organisms (in general) in various North American freshwater ecosystems	Compare sediment chemical concentrations to TELs	Sediment	Yes. Reliable predictor of Level 0 and 1 hits (i.e., no or unlikely adverse effects). Reasonable to use to identify locations "rarely associated with adverse biological effects".	High false positive rate, so unreliable for predicting Level 2 or 3 hits (likely adverse effects).	No; narrative intent inconsistent with assessment endpoint	Yes--relatively consistent with TECs and ER-Ls	No. TELs are screening-level thresholds for identify locations "rarely associated with adverse biological effects" (broadly defined) on aquatic organisms (in general) in various North American freshwater ecosystems. As such they're not appropriate for deriving Portland Harbor PRGs. Instead, they should have been used to delineate portions of the Study Area where bioassays were unnecessary (i.e., being below the TELs should have screened areas out of the benthic BERA).

LOE	Method for Assessing LOE	Assessment Tools	Assessment(s)	Potential PRG Matrix	Strong Line? Why?	What issues affect reliability and certainty of LOE?	Can a Numeric PRG be Derived? (i.e., evidence of a relationship to sediment)	Corroborated by other LOEs? If so, which ones and what are their strengths.	LWG Position - Should this LOE be used to derive PRGs for use in FS?
		2. **Probable effect levels (PELs)** for predicting freshwater sediment concentrations "frequently associated with adverse biological effects" (broadly defined) on aquatic organisms (in general) in various North American freshwater ecosystems	Compare sediment chemical concentrations to PELs Compare mean PEL quotient to a mean quotient threshold of 0.7 (the 0.7 threshold value was used as directed by EPA in the draft BERA problem formulation)		No; PELs don't predict empirical bioassay results	Unreliable for predicting empirical bioassay results	Yes, but unreliable	No	No. PELs could have been used to derive PRGs had a site-specific benthic toxicity and chemistry study not been conducted, but it was, and so site-specific SQGs should be used rather than generic values. Moreover, the PELs were found to be unreliable predictors of the empirical bioassay results (hit classifications) for the site.
		3. **Threshold effect concentrations (TECs)** for predicting freshwater sediment concentrations below which **none** of the following effects are expected to occur: a) amphipod (*Hyalella azteca*) survival, growth or reproduction significantly different from the responses observed in reference or control sediments b) mayfly (*Hexagenia limbata*) survival or growth significantly different from the responses observed in reference or control sediments c) midge *(Chironomus dilutus* or *Chironomus riparius*) survival, growth or deformities significantly different from the responses observed in reference or control sediments d) oligochaete (*Lumbriculus variegatus*) survival significantly different from the responses observed in reference or control sediments e) daphnid (*Cereodaphnia dubia*) survival significantly different from the responses observed in reference or control sediments AND f) bacterial (*Photobacterium phosphoreum*) luminescence (i.e., Microtox) significantly different from the responses observed in reference or control sediments	Compare sediment chemical concentrations to TECs		Yes. Reliable predictor of Level 0 and 1 hits (i.e., no or unlikely adverse effects). Reasonable to use for predicting freshwater sediment concentrations below which adverse effects are not expected to occur.	High false positive rate, so unreliable for predicting Level 2 or 3 hits (likely adverse effects).	No; narrative intent inconsistent with assessment endpoint	Yes--relatively consistent with TELs and ER-Ls	No. TECs are screening-level thresholds for identifying locations where adverse effects are not expected to occur.

LOE	Method for Assessing LOE	Assessment Tools	Assessment(s)	Potential PRG Matrix	Strong Line? Why?	What issues affect reliability and certainty of LOE?	Can a Numeric PRG be Derived? (i.e., evidence of a relationship to sediment)	Corroborated by other LOEs? If so, which ones and what are their strengths.	LWG Position - Should this LOE be used to derive PRGs for use in FS?
		4. **Probable effect concentrations (PECs)** for predicting freshwater sediment concentrations above which **any one or more** of the following effects is expected to occur "more often than not:" a) amphipod (*Hyalella azteca*) survival, growth or reproduction significantly different from the responses observed in reference or control sediments b) mayfly (*Hexagenia limbata*) survival or growth significantly different from the responses observed in reference or control sediments c) midge (*Chironomus dilutus* or *Chironomus riparius*) survival, growth or deformities significantly different from the responses observed in reference or control sediments d) oligochaete (*Lumbriculus variegatus*) survival significantly different from the responses observed in reference or control sediments e) daphnid (*Cereodaphnia dubia*) survival significantly different from the responses observed in reference or control sediments OR f) bacterial (*Photobacterium phosphoreum*) luminescence (i.e., Microtox) significantly different from the responses observed in reference or control sediments	Compare sediment chemical concentrations to PECs Compare mean PEC quotient to a mean quotient threshold of 0.7 (the 0.7 threshold value was used as directed by EPA in the draft BERA problem formulation)		No; PECs don't predict empirical bioassay results	Unreliable for predicting empirical bioassay results	Yes, but unreliable	No	No. PECs could have been used to derive PRGs had a site-specific benthic toxicity and chemistry study not been conducted, but it was, and so site-specific SQGs should be used rather than generic values. Moreover, the PECs were found to be unreliable predictors of the empirical bioassay results (hit classifications) for the site.
		5. **Washington State sediment quality standards (SQS)** for predicting marine sediment quality that will result in no adverse effects, including no acute or chronic adverse effects on biological resources and no significant risk to human health.	Compare sediment chemical concentrations to SQS		No; don't predict empirical bioassay results	Unreliable for predicting empirical bioassay results	No; narrative intent inconsistent with assessment endpoint	No	No. SQS are regulatory thresholds (in Washington state) for identifying locations in marine sediment where no adverse effects to biological resources are likely to occur. As such they're not appropriate for deriving Portland Harbor PRGs. Because they're marine sediment thresholds, they shouldn't be used to delineate portions of the Study Area where bioassays were unnecessary (i.e., freshwater screening values are available and should have been used instead of, not in addition to, the marine SQS).
		6. **Washington State sediment cleanup screening levels (CSLs)** for identifying Puget Sound sediment cleanup sites per WAC 173-204-530 procedures.	Compare sediment chemical concentrations to CSLs Compare mean CSL quotient to a mean quotient threshold of 0.7 (the 0.7 threshold value was used as directed by EPA in the draft BERA problem formulation)		No; don't predict empirical bioassay results	Unreliable for predicting empirical bioassay results	Yes, but unreliable	No	No. The CSLs are the minimum cleanup levels to be applied to marine sediment where both chemical and biological standards have been exceeded. Biological (and chemical) testing has been done for Portland Harbor; that alone is reason not to use the CSLs as PRGs. Site-specific SQGs should be used rather than generic values. Moreover, the CSLs were found to be unreliable predictors of the empirical bioassay results (hit classifications) for the site.

LOE	Method for Assessing LOE	Assessment Tools	Assessment(s)	Potential PRG Matrix	Strong Line? Why?	What issues affect reliability and certainty of LOE?	Can a Numeric PRG be Derived? (i.e., evidence of a relationship to sediment)	Corroborated by other LOEs? If so, which ones and what are their strengths.	LWG Position - Should this LOE be used to derive PRGs for use in FS?
		7. **Effect range-low (ER-L)** for predicting marine and estuarine sediment concentrations below which none of the following adverse effects is expected to occur except "rarely": a) depressed species richness b) low total abundance c) "significantly" or "relatively" elevated sediment toxicity (test species not specified) d) histopathological disorders in demersal fish observed in field studies e) spiked sediment single chemical laboratory bioassay EC50 or LC50 concentration exceeded AND f) toxicity predicted by equilibrium partitioning theory The ER-L is defined as the lower 10th percentile of the authors' compiled adverse effects dataset.	Compare sediment chemical concentrations to ER-Ls		Yes. Reliable predictor of Level 0 and 1 hits (i.e., no or unlikely adverse effects). Reasonable to use to identify locations where adverse effects is expected to rarely occur.	High false positive rate, so unreliable for predicting Level 2 or 3 hits (likely adverse effects).	No; narrative intent inconsistent with assessment endpoint	Yes--relatively consistent with TECs and TELs	No. ER-Ls are screening-level thresholds for identifying locations (in marine and estuarine sediment) where adverse effects is expected to rarely occur.
		8. **Effect range-median (ER-M)** for predicting marine and estuarine sediment concentrations above which **any one or more** of the following adverse effects is expected to occur "frequently." a) depressed species richness b) low total abundance c) "significantly" or "relatively" elevated sediment toxicity (test species not specified) d) histopathological disorders in demersal fish observed in field studies e) spiked sediment single chemical laboratory bioassay EC50 or LC50 concentration exceeded AND f) toxicity predicted by equilibrium partitioning theory The ER-M is defined as the median (50th percentile) of the authors' compiled adverse effects dataset.	Compare sediment chemical concentrations to ER-Ms Compare mean ER-M quotient to a mean quotient threshold of 0.7 (the 0.7 threshold value was used as directed by EPA in the draft BERA problem formulation)		No; don't predict empirical bioassay results	Unreliable for predicting empirical bioassay results	Yes, but unreliable	No	No. ER-Ms could have been used to derive PRGs had a site-specific benthic toxicity and chemistry study not been conducted, but it was, and so site-specific SQGs should be used rather than generic values. Moreover, the ER-Ms were found to be unreliable predictors of the empirical bioassay results (hit classifications) for the site. Finally, ER-Ms are marine and estuarine sediment screening levels. Freshwater screening values are available and should have been used instead of, not in addition to, the marine and estuarine sediment screening values

LOE	Method for Assessing LOE	Assessment Tools	Assessment(s)	Potential PRG Matrix	Strong Line? Why?	What issues affect reliability and certainty of LOE?	Can a Numeric PRG be Derived? (i.e., evidence of a relationship to sediment)	Corroborated by other LOEs? If so, which ones and what are their strengths.	LWG Position - Should this LOE be used to derive PRGs for use in FS?
		9. **Regional Sediment Evaluation Team (RSET) SL1** (interim freshwater lower screening level) values. SL1 values are concentrations below which adverse effects to benthic organisms are not expected.	Compare sediment chemical concentrations to SL1 values		No; these lower levels don't predict empirical bioassay results	Unreliable for predicting empirical bioassay results	No; narrative intent inconsistent with assessment endpoint	No	No. SL1 values are screening-level thresholds for identifying locations where adverse effects to benthic organisms are not expected. As such they're not appropriate for deriving Portland Harbor PRGs. Instead, they should have been used to delineate portions of the Study Area where bioassays were unnecessary (i.e., being below the SL1s should have screened areas out of the benthic BERA).
		10. **RSET SL2** (interim freshwater upper screening level) values. SL2 values are concentrations at which minor adverse effects might be observed in the more sensitive groups of benthic organisms.	Compare sediment chemical concentrations to SL2 values Compare mean SL2 quotient to a mean quotient threshold of 0.7 (the 0.7 threshold value was used as directed by EPA in the draft BERA problem formulation)		No; these upper screening levels don't predict empirical bioassay results	Unreliable for predicting empirical bioassay results	Yes, but unreliable	No	No. SL2 values could have been used to derive PRGs had a site-specific benthic toxicity and chemistry study not been conducted, but it was, and so site-specific SQGs should be used rather than generic values. Moreover, the SL2s were found to be unreliable predictors of the empirical bioassay results (hit classifications) for the site.
		11. **Equilibrium partitioning sediment benchmarks (ESBs)** for PAH mixtures, non-ionic organic compounds, gamma hexachlorocyclohexane (HCH), endrin and dieldrin. ESB is derived by multiplying a chemical's water-quality based final chronic value (FCV) or species chronic value (SCV) by its Koc, yielding the concentration in sediment that should provide the same level of protection that the FCV or SCV provides in water, assuming equilibrium between sediment and the water to which organisms are exposed. ESB should be interpreted as a chemical concentration below which adverse effects are not expected and above which adverse effects might occur.	Compare sediment chemical concentrations to ESB values		No; don't predict empirical bioassay results	Unreliable for predicting empirical bioassay results	Yes, but unreliable	No	No. ESB values could have been used to derive PRGs had a site-specific benthic toxicity and chemistry study not been conducted, but it was, and so site-specific SQGs should be used rather than generic values. Moreover, the ESB values were found to be unreliable predictors of the empirical bioassay results (hit classifications) for the site.
Effect of laboratory exposure to Study Area sediment on *Corbicula fluminea* biomass	28-day *C. fluminea* bioaccumulation test	Perform a qualitative toxicity assessment based on the growth measured during bioaccumulation tests as directed by EPA in the draft BERA problem formulation. Growth estimates were calculated as final biomass divided by the initial estimated control biomass. The estimated growth in test sediments was then compared to estimated growth in the negative control group. Differences could not be statistically tested due to lack of replication.	Sediment	No. Bioaccumulation tests not designed for this purpose	Qualitative only.	No	Yes; empirical bioassays that are directly linked to actual site conditions/effects	No. Uncertainty about the bioaccumulation test biomass data is unquantifiable. The data should not be used to derive PRGs. The evidence should be considered corroborative only.

LOE	Method for Assessing LOE	Assessment Tools	Assessment(s)	Potential PRG Matrix	Strong Line? Why?	What issues affect reliability and certainty of LOE?	Can a Numeric PRG be Derived? (i.e., evidence of a relationship to sediment)	Corroborated by other LOEs? If so, which ones and what are their strengths.	LWG Position - Should this LOE be used to derive PRGs for use in FS?
	Effect of laboratory exposure to Study Area sediment on *Corbicula fluminea* survival	28-day *C. fluminea* bioaccumulation test	Perform a qualitative toxicity assessment based on the survival measured during bioaccumulation tests as directed by EPA in the draft BERA problem formulation. To assess the suitability of using *Chironomus* and *Hyalella* survival test results as a surrogate for clams, 10-day *Chironomus dilutus* and 28-day *Hyalella azteca* survival test results were compared with *C. fluminea* survival results measured as part of the bioaccumulation tests, using "nearest neighbor" station comparisons.	Sediment	No. Bioaccumulation tests not designed for this purpose.	Qualitative only.	No	Yes; empirical bioassays that are directly linked to actual site conditions/effects	No. Uncertainty about the bioaccumulation test survival data is unquantifiable. The data should not be used to derive PRGs. The evidence should be considered corroborative only.
	Effect of field exposure to Study Area sediment and water on chemical concentrations in clam tissue	Field-collected *C. fluminea* tissue	Compare field-collected tissue residue concentrations to tissue TRVs	Sediment	Yes; field collected tissue likely represent steady-state conditions and actual bioavailability of sediment- and water-borne contaminants	Questions about appropriateness of tissue TRVs (e.g., chronic toxicity estimated from acute data by applying an ACR, a practice which not all ecotoxicologists find technically defensible)	Yes, for selected chemicals. Tissue TRVs were developed for 10 benthic COPCs (As, Cd, Cu, Zn, TBT, BEHP, dibutyl phthalate, total PCBs, DDD and total DDx). A relationship between clam tissue and sediment concentrations was found for only two of the 10 (total PCBs and total DDx) so only two clam tissue-based PRGs can be derived. None of the 41 field clam samples exceeded the tissue TRV for total DDx, and only one tissue samples (2.4%) exceeded the tissue TRV for total PCBs.	Yes, in that it is not inconsistent with other LOEs	Yes, for PCBs and total DDx.
	Effect of field exposure to Study Area sediment and water on chemical concentrations in mussel tissue	Field-collected mussel (*Margaritifera falcata* and *Anodonta nuttalliana*) tissue	Compare field-collected tissue residue concentrations to tissue TRVs	Sediment	Yes; field collected tissue likely represent steady-state conditions and actual bioavailability of sediment- and water-borne contaminants	The final species LOAEL for mussels (*Dreissena polymorpha*) was 69 mg/kg ww. The maximum measured tissue concentration in a mussel sample was 41 mg/kg ww. The tissue TRV was 24.07 mg/kg ww, driven by amphipod and snail mortality. Additionally, few mussels were collected in the LWR limiting the spatial representation of this LOE.	The only tissue TRV exceedances in mussels were for Zn (maximum HQ = 1.7). There's no relationship between Zn tissue and sediment concentrations.	Yes, in that it is not inconsistent with other LOEs	No; no relationship between tissue and sediment Zn concentrations.

LOE	Method for Assessing LOE	Assessment Tools	Assessment(s)	Potential PRG Matrix	Strong Line? Why?	What issues affect reliability and certainty of LOE?	Can a Numeric PRG be Derived? (i.e., evidence of a relationship to sediment)	Corroborated by other LOEs? If so, which ones and what are their strengths.	LWG Position - Should this LOE be used to derive PRGs for use in FS?
	Effect of field exposure to Study Area sediment, water and prey tissue on chemical concentrations in crayfish tissue	Field-collected crayfish (*Pacifastacus leniusculus*) tissue	Compare field-collected tissue residue concentrations to tissue TRVs	Sediment	Yes; field collected tissue likely represent steady-state conditions and actual bioavailability of sediment- and water-borne contaminants and prey.	Interspecies extrapolation issue. Tissue copper concentrations were nearly constant across the 32 crayfish samples regardless of sediment concentrations, and they all had HQs > 1 (max HQ = 2.6).	No. No relationship between tissue and sediment concentrations.	Yes, in that it is not inconsistent with other LOEs	No. No relationship between crayfish tissue and sediment Cu concentrations.
	Effect of laboratory exposure to Study Area sediment on chemical concentrations in clam tissue	Steady state-adjusted tissue residue estimates based on 28-day clam (*C. fluminea*) bioaccumulation tests	Compare steady state-adjusted laboratory-exposed tissue residue concentrations to tissue TRVs	Sediment	No due to lab to field extrapolation and steady-state extrapolation	Questions about appropriateness of tissue TRVs, about steady state adjustment, about exposure regime. Why use when we have field data? No (0/41) field clams had tissue residues exceeding the total DDT TRV, and only one lab clam sample had HQs > 1 for three chemicals (max HQ 2.2).	One lab clam exceeded the tissue TRVs for TBT, BEHP and total DDx. A relationship between clam tissue and sediment concentrations was found for total DDx but not for TBT or BEHP.	Yes, consistent with the field clam LOE in indicating little or no risk to clams	No. Weak LOE, should rely on the field data.
	Effect of laboratory exposure to Study Area sediment on chemical concentrations in oligochaete worm tissue	Steady state-adjusted tissue residue estimates based on 28-day oligochaete worm (*Lumbriculus variegatus*) bioaccumulation tests	Compare steady state-adjusted laboratory-exposed tissue residue concentrations to tissue TRVs	Sediment	No due to lab to field extrapolation and steady-state extrapolation	Questions about appropriateness of tissue TRVs, about steady state adjustment, and about lab to field extrapolation	Lab worm tissue residues exceeded tissue TRVs for As (2/35, max HQ = 1.5), Cu (1/35, max HQ = 2.6), Zn (27/35, max HQ = 1.3), TBT (1/35, max HQ = 11), total PCBs (1/35, max HQ = 1.2) and total DDx (2/35, max HQ = 3.2). Of these a relationship between worm tissue and sediment concentrations was found for TBT, total PCBs and total DDx.	Yes, in that it is not inconsistent with other LOEs	Yes for TBT, PCBs, and total DDx (i.e., chemicals demonstrating a relationship between sediment and tissue). Also, these are the only data we have for worm tissue (no equivalent field data).
	Predicted effect of field exposure to Study Area sediment and water on chemical concentrations in clam tissue	Bioaccumulation model-predicted clam tissue concentrations in locations where clams weren't collected	Compare tissue-residue concentrations estimated using a bioaccumulation model to tissue TRVs	Sediment	Yes for selected chemicals (PCBs, TEQs, and DDx), species predictive accuracy factors (SPAF) were generally <5 for invertebrate species evaluated.	Modeled tissue concentrations less representative than field collected clam tissue concentrations. Empirical data did not exceed tissue TRVs except in one sample for total PCBs; model tended to underpredict tissue residues for this chemical group.	Yes	Yes, in that it is not inconsistent with other LOEs	No. Benthic toxicity is a stronger line of evidence for the development of benthic PRGs.

LOE	Method for Assessing LOE	Assessment Tools	Assessment(s)	Potential PRG Matrix	Strong Line? Why?	What issues affect reliability and certainty of LOE?	Can a Numeric PRG be Derived? (i.e., evidence of a relationship to sediment)	Corroborated by other LOEs? If so, which ones and what are their strengths.	LWG Position - Should this LOE be used to derive PRGs for use in FS?
	Predicted effect of field exposure to Study Area sediment, water and prey tissue on chemical concentrations in crayfish tissue	Bioaccumulation model-predicted crayfish tissue concentrations in locations where crayfish weren't collected	Compare tissue-residue concentrations estimated using a bioaccumulation model to tissue TRVs	Sediment	Yes for selected chemicals (PCBs, TEQs, and DDx), species predictive accuracy factors (SPAF) were generally <5 for benthic species evaluated.	Modeled tissue concentrations less representative than field collected crayfish tissue concentrations; model typically overpredicted tissue residues. Empirical data did not exceed TRVs for bioaccumulative chemicals.	Yes	Yes, in that it is not inconsistent with other LOEs	No. Benthic toxicity is a stronger line of evidence for the development of benthic PRGs.
	Predicted effect of field exposure to Study Area sediment and water on chemical concentrations in oligochaete worm tissue	Bioaccumulation model-predicted oligochaete worm tissue concentrations	Compare tissue-residue concentrations estimated using a bioaccumulation model to tissue TRVs	Sediment	Yes for selected chemicals (PCBs, TEQs, and DDx), species predictive accuracy factors (SPAF) were generally <5 for benthic species evaluated.	Tissue concentrations used in model based on steady-state predictions; unlikely to be representative of field conditions (few empirical laboratory samples exceeded bioaccumulative chemical TRVs). Model typically overpredicted tissue residues for this trophic level.	Yes	Yes, in that it is not inconsistent with other LOEs	No. Benthic toxicity is a stronger line of evidence for the development of benthic PRGs.
Toxicity of chemicals in surface water to benthic invertebrates	Effect of field exposure to Study Area sediment, water and prey tissue on chemical concentrations in epibenthic invertebrate tissue	Field-collected epibenthic invertebrate tissue (multiplate samples)	Compare field-collected tissue residue concentrations to tissue TRVs	Surface water	No; field collected organisms were not in direct contact with sediment	Limited biomass; not associated with sediment	No	Yes, in that it is not inconsistent with other LOEs	No. No tissue TRV exceedances.
	Predicted effects to benthic organisms based on a comparison of water chemical concentrations to TRVs	Surface water chemical concentrations (all sampling methods) EPA-approved water TRVs	Compare detected concentrations in individual surface water samples to water TRVs.	Surface water	No; exposure not directly linked to sediment. Comparison provides a screening-level assessment only.	The source of surface water chemicals is unlikely to be proximal sediment, which is a major source of uncertainty in a study of potential sediment remedies.	Yes	Yes, in that it is not inconsistent with other LOEs	No; unless applied very carefully surface water PRGs can be misleading in a feasibility study of potential sediment remedies.
Toxicity of chemicals in shallow transition zone water (TZW) to benthic invertebrates	Predicted effects based on a comparison of shallow TZW chemical concentrations to TRVs	TZW chemical concentrations (all sampling methods) EPA-approved water TRVs	Compare detected concentrations in individual shallow TZW samples to water TRVs.	Shallow TZW	No; the draft BERA only provides a screening-level exposure assessment	The analysis presented in the draft BERA ignores relevant and appropriate information about ecological exposure to TZW	Yes	Yes; consistent with the sediment LOEs	No. Use of the TRVs provide only a screening level assessment and are not appropriate for use as risk-based PRGs in this context. At least not unless EPCs are better estimated than we were allowed to do in the draft BERA.
Benthic community structure	Observed benthic community successional stages	Sediment profile imagery (SPI) survey data collected throughout the Study Area	Compare SPI data to expectations based on physical characteristics of survey stations	n/ap (corroborative LOE)	Yes, images allow assessment of the maturity of the benthic community and interpretation of physical vs other stressors	Qualitative LOE; SPI images do not provide a quantitative assessment of the survival, growth or reproduction in benthic organisms nor were they co-located with sediment chemistry and toxicity such that a quantitative link could be developed.	No	Yes; consistent with the other sediment LOEs	n/ap (corroborative LOE)

LOE	Assessment Tools	Assessment(s)	Potential PRG Matrix	Strong Line? Why? What issues affect reliability and certainty of LOE?	Can a Numeric PRG be Derived? (i.e., evidence of a relationship to sediment)	Corroborated by other LOEs? If so, which ones and what are their strengths.	LWG Position - Should this LOE be used to derive PRGs for use in FS?
Chemical body burden	Field-collected largescale sucker tissue samples, fish tissue residue TRVs Field-collected juvenile white sturgeon tissue samples, fish tissue residue TRVs Field-collected juvenile Chinook salmon tissue samples, fish tissue residue TRVs Field-collected peamouth tissue samples, fish tissue residue TRVs Field-collected sculpin tissue samples, fish tissue residue TRVs Field-collected smallmouth bass tissue samples, fish tissue residue TRVs Field-collected northern pikeminnow tissue samples, fish tissue residue TRVs Field-collected Pacific lamprey tissue samples, fish tissue residue TRVs	Compare measured body burdens to literature-based tissue residue TRVs	Sediment	TRVs are not species-specific. Strength of the evidence varies by chemical depending on mode and site of toxicity, how the chemical's distributed in tissues and whether and how the organisms bioregulates the chemical. Strongest for PBTs. Lack of consensus among ecotoxicologists about appropriateness and relevance of whole body tissue residues as an indicator of ecotoxicity for metals and bioregulated organics. Exposure and effect uncertainties for all fish tissue residue COCs are summarized in draft BERA Table 7-13. Risks to sculpin were determined based on an evaluation of each individual composite fish tissue sample. This approach was used as directed by EPA in order to identify locations within the Study Area where adverse effects might occur to fish within the populations that are being assessed. Sample-by-sample assessment is a conservative and questionable method for evaluating risks to populations. It relies on inferences that have little or no scientific basis because population-level processes compensate for adverse effects on individuals (Pastorok *et al.* 2001). Sample-level evaluations do not represent population-level effects. Several methods have been used elsewhere in an attempt to address potential population-level effects but no consensus approach currently exists; other than HQ approaches do not yield population-level assessments.	The only fish tissue residue COCs for which numeric PRGs can be developed are total PCBs, total DDx and Pb.	Focusing on the COCs for which numeric PRGs can be developed: total PCBs and total DDx are identified as COCs based on the surface water LOE but not based on the dietary dose LOE.	Yes for total PCBs and total DDx. No for Pb because the lead BSAF is only significant due to the high statistical influence of an outlier.
Dietary dose	Estimated dietary doses for largescale sucker, dietary dose TRVs Estimated dietary doses for juvenile white sturgeon, dietary dose TRVs Estimated dietary doses for juvenile Chinook salmon, dietary dose TRVs Estimated dietary doses for peamouth, dietary dose TRVs Estimated dietary doses for sculpin, dietary dose TRVs Estimated dietary doses for smallmouth bass, dietary dose TRVs Estimated dietary doses for northern pikeminnow, dietary dose TRVs Estimated dietary doses for Pacific lamprey, dietary dose TRVs	Compare estimated dietary doses to literature-based dietary dose TRVs	Sediment	TRVs are not species-specific. Strength of the evidence varies by chemical. The only fish dietary dose COC for which a numeric PRG can be developed is TBT, and that requires relying on a BSAR/F for lab worms. The uncertainties associated with using lab bioaccumulation testing to represent prey chemical concentrations were summarized on p. 282 of the draft BERA. There's also uncertainty associated with the lack of pelagic prey (see draft BERA, p. 283). The TBT TRV is highly uncertain for reasons summarized on p. 285 of the draft BERA. Also, for fish, the dietary dose approach is not commonly used in ecological risk assessment and limited data are available to calculate dietary dose TRVs. Table 7-33 of the draft BERA summarizes uncertainty about dietary prey portions. Table 7-32 summarizes exposure and effect uncertainties for all fish dietary COCs. Reliance on HQs in the BERA is itself a reliability limitation (e.g., see Allard, Fairbrother, Hope, Hull, Johnson, Kaputska, Mann, McDonald and Sample 2010. Recommendations for the Development and Application of Wildlife Toxicity Reference Values. *Integrated Environmental Assessment and Management* **6**(1):28-37.).	The only fish dietary dose COC for which numeric PRGs can be developed is TBT.	Focusing on the COC for which numeric PRGs can be developed: TBT is not identified as a COC by the tissue residue LOE and HQ exceedances are .	No. The uncertainties about TBT exposure, dietary dose TRV and BSAR/F are all high and the LOE is uncorroborated.
Surface water	Surface water concentrations, water TRVs	Compare measured surface water concentration to literature-based surface water TRVs	Water	The strength of the LOE varies by chemical. For example the TRVs for benzo(a)anthracene and benzo(a)pyrene were based on invertebrate (*Daphnia* sp.) toxicity data and the TRV for DDTs is based on the 4,4'-DDT AWQC, which represents effects on brown pelican, so this LOE is particularly weak for those COCs.	Yes (there's no extrapolation across media so the TRVs could be used as water PRGs).	Aluminum, zinc, benzo(a)anthracene, benzo(a)pyrene, naphthalene, BEHP and TCE were uncorroborated by other LOEs.	Yes, if EPA surface water PRGs are to be used then this LOE should be used to derive PRGs for the draft FS.
Transition zone water	TZW concentrations, water TRVs	Compare measured TZW concentration to literature-based water TRVs	TZW	This is a weak LOE because EPA directed the LWG to screen undiluted TZW against water TRVs. And prohibited the LWG from deriving baseline-quality EPCs.	Yes (there's no extrapolation across media so the TRVs could be used as water PRGs, or the TRVs could be adjusted to more accurately account for fish exposure to TZW).	Some of the TZW chemicals that exceeded water TRVs were also elevated in sediment or tissues.	No, but the chemicals should be carried into the FS and the evaluation of remedial action alternatives in the draft FS should assess whether the threshold criterion of protectiveness is met.
Benthic fish health and PAH exposure	Fish health field observations (occurrence of lesions and abnormalities), literature-based sediment concentrations associated with lesion occurrence	Compare prevalence of lesions and abnormalities in the Study Area and the Lower Columbia River	Sediment	This is a very weak LOE. The results were inconclusive concerning any possible relationship between PAH exposure and incidence of lesions. Population-level effects could not be extrapolated from an organism-level effect	No	The LOE was inconclusive	No

LOE	Assessment Tools	Assessment(s)	Potential PRG Matrix	Strong Line? Why?	What issues affect reliability and certainty of LOE?	Can a Numeric PRG be Derived? (i.e., evidence of a relationship to sediment)	Corroborated by other LOEs? If so, which ones and what are their strengths.	LWG Position - Should this LOE be used to derive PRGs for use in FS?
Dietary dose	EPCs from dietary dose estimates based on sediment, field clam and steady state-adjusted lab worm and clam data; dietary dose TRVs (TRVs not species specific)	Compare estimated dietary doses to literature-based dietary dose TRVs	Sediment	The only LOE for spotted sandpiper. TRVs are not species-specific. Strength of the evidence varies by chemical. The uncertainty associated with basing dietary exposure assumptions on steady-state adjusted bioaccumulation test data are summarized on p. 385 of the draft BERA. The B(a)P PRG requires relying on a weak (r^2 = 0.39) lab worm BSAR (no relationship between clam tissue and sediment B(a)P concentrations . Other PRGs are based on the FWM. Key uncertainties in the bird dietary dose TRVs are summarized in Table 8-11 of the draft BERA. Reliance on HQs in the BERA is itself a reliability limitation (see, e.g., Allard *et al.* (2010).		Numeric PRGs can be developed for B(a)P, total PCBs, PCB TEQ, dioxin/furan TEQ, aldrin and total DDx. The LOE is weakest for B(a)P.	No other lines	Yes

LOE	Assessment Tools	Assessment(s)	Potential PRG Matrix	Strong Line? Why?	What issues affect reliability and certainty of LOE?	Can a Numeric PRG be Derived? (i.e., evidence of a relationship to sediment)	Corroborated by other LOEs? If so, which ones and what are their strengths.	LWG Position - Should this LOE be used to derive PRGs for use in FS?
Dietary dose	EPCs from dietary dose estimates based on peamouth, sculpin, crayfish and field clam and steady state-adjusted lab clam data, dietary dose TRVs (TRVs not species specific)	Compare estimated dietary doses to literature-based dietary dose TRVs	Sediment	The only LOE for hooded merganser. TRV is not species-specific. It is possible that the PCB TRV for hooded merganser has been overestimated by an order of magnitude. The mallard LOAEL is approximately 30x higher than the chicken chick LOAEL (based on reduced hatchability) that was used for the TRV. (This is significant given that the maximum HQ was only 3.8). See also draft BERA Table 8-23 for an uncertainty summary. Reliance on HQs in the BERA is itself a reliability limitation.		The only COC is total PCBs. A numeric PRG can be developed.	No other lines	Yes

LOE	Assessment Tools	Assessment(s)	Potential PRG Matrix	Strong LOE? Why?	What issues affect reliability and certainty of LOE?	Can a Numeric PRG be Derived? (i.e., evidence of a relationship to sediment)	Corroborated by other LOEs? If so, which ones and what are their strengths.	LWG Position - Should this LOE be used to derive PRGs for use in FS?
Dietary dose	EPCs from dietary dose estimates based on peamouth, sculpin, crayfish and field clam and steady state-adjusted lab clam data, dietary dose TRVs (TRVs not species specific)	Compare estimated dietary doses to dietary dose TRVs	Sediment	TRVs are not species-specific. Weak LOE for lead. The LOAEL is based on an EcoSSL that's an order of magnitude lower than the lowest acceptable literature-based LOAEL. Smallmouth bass contributed 100% of risk, based on a single outlier that's more than 100x greater than the other smallmouth bass concentration available from the same exposure areas and 2-5 orders of magnitude higher than lead concentrations detected in all other smallmouth bass samples. Strongest available LOE for total PCBs; see draft BERA Table 8-31 for an uncertainty summary.		The COCs by this LOE are lead and total PCBs.	Lead--no, total PCBs--yes	Yes for total PCBs, no for lead
Bird egg tissue residue	Bird egg concentrations predicted by multiplying estimated prey tissue concentrations by estimated prey-to-egg tissue BMFs from the literature; bird egg TRVs from the literature	Compare estimated bird egg concentrations to TRVs	Sediment	No. The BMFs are highly unreliable. See draft BERA pp. 381 and 383 for a summary of BMF uncertainties.		The COCs are total PCBs, PCB TEQ, PCDD/F TEQ, total TEQ and sum DDE. PRGs can be derived for all except total TEQ.	Total PCBs--yes (qualitatively). All other COCs--no.	No

LOE	Assessment Tools	Assessment(s)	Potential PRG Matrix	Strong Line? Why?	What issues affect reliability and certainty of LOE?	Can a Numeric PRG be Derived? (i.e., evidence of a relationship to sediment)	Corroborated by other LOEs? If so, which ones and what are their strengths.	LWG Position - Should this LOE be used to derive PRGs for use in FS?
Dietary dose	EPCs from dietary dose estimates based on peamouth, sculpin, crayfish and field clam and steady state-adjusted lab clam data, dietary dose NOAEL TRVs	Compare estimated dietary doses to NOAELTRVs	Sediment	Strongest available LOE	TRVs are not species-specific. See draft BERA pp. 419 and 420 for a summary of uncertainties and Table 8-28	The COCs by this LOE are mercury and total PCBs. A numeric PRG can be derived for total PCBs but not for mercury.	Mercury--no, total PCBs--yes	Yes for total PCBs
Bird egg tissue residue	Bird egg concentrations predicted by multiplying estimated prey tissue concentrations by estimated prey-to-egg tissue BMFs from the literature; NOAEL bird egg TRVs from the literature	Compare estimated bird egg concentrations to NOAAEL TRVs	Sediment	No	The BMFs are highly unreliable. See draft BERA pp. 381 and 383 for a summary of BMF uncertainties.	The COCs are total PCBs, PCB TEQ, PCDD/F TEQ, total TEQ and 4,4' DDE. PRGs can be derived for all except total TEQ.	Total PCBs--yes (qualitatively). All other COCs--no.	No

LOE	Assessment Tools	Assessment(s)	Potential PRG Matrix	Strong Line? Why?	What issues affect reliability and certainty of LOE?	Can a Numeric PRG be Derived? (i.e., evidence of a relationship to sediment)	Corroborated by other LOEs? If so, which ones and what are their strengths.	LWG Position - Should this LOE be used to derive PRGs for use in FS?
Dietary dose	EPCs from dietary dose estimates based on crayfish, largescale sucker, carp, sculpin, smallmouth bass and clam (field and steady state-adjusted lab) data, dietary dose TRVs	Compare estimated dietary doses to dietary dose TRVs	Sediment	It's the only LOE	See Table 8-35 for a summary of TRV and exposure uncertainties	The COCs are total PCBs and total TEQ. A numeric PRG can be derived for total PCBs but not for total TEQ.	n/ap	Yes, for total PCBs

LOE	Assessment Tools	Assessment(s)	Potential PRG Matrix	Strong Line? Why?	What issues affect reliability and certainty of LOE?	Can a Numeric PRG be Derived? (i.e., evidence of a relationship to sediment)	Corroborated by other LOEs? If so, which ones and what are their strengths.	LWG Position - Should this LOE be used to derive PRGs for use in FS?
Dietary dose	EPCs from dietary dose estimates based on crayfish, largescale sucker, carp, sculpin and smallmouth bass data, dietary dose TRVs	Compare estimated dietary doses to dietary dose TRVs	Sediment	It's the only LOE, and it's the strongest LOE in the draft BERA because we have a species-specific TRV for the risk driver (total PCBs) and strong data with which to assess mink exposure.	See Table 8-33 for a summary of TRV and exposure uncertainties. The lead risk estimate is due exclusively to the smallmouth bass outlier from RM 9.5-RM 10.5. Smallmouth bass prey contributed 100% of risk, based on that single outlier that's more than 100x greater than the other smallmouth bass concentration available from the same exposure area and 2-5 orders of magnitude higher than lead concentrations detected in all other smallmouth bass samples.	The COCs are Pb, total PCBs and total TEQ. Numeric PRGs can be derived for Pb and total PCBs but not for total TEQ.	n/ap	Yes, for total PCBs. No for Pb.